



TITLE:

# 医薬品-標的タンパク質-副作用の関連解析による副作用の発現機序推定( Abstract\_要旨 )

AUTHOR(S):

水谷, 紗弥佳

---

CITATION:

水谷, 紗弥佳. 医薬品-標的タンパク質-副作用の関連解析による副作用の発現機序推定. 京都大学, 2015, 博士(理学)

ISSUE DATE:

2015-03-23

URL:

<https://doi.org/10.14989/doctor.k18839>

RIGHT:

京都大学	博士（理 学）	氏名	水谷 紗弥佳
論文題目	医薬品-標的タンパク質-副作用の関連解析による副作用の発現機序推定		
(論文内容の要旨)			
<p>医薬品としての薬物の副作用は、医療の安全性を考える上で非常に重要な問題である。今までの研究では、薬物の化学構造に基づき副作用を予測する方法が開発されてきた。しかしながら、副作用には様々な発現パターンがあり、安全な薬物治療のための副作用の回避にはその作用機序を見極める必要がある。また、薬理作用の影響は、標的分子への局所的な活性制御に留まらず、生化学経路、細胞、組織、臓器など複雑な階層構造を通じて生体全体へ波及する。そのため、副作用の発現機序の推定や予測には、生体システム全体に関する情報を統合活用することが必要である。さらに、副作用は患者の個人差によるところもあり、大規模で多様な集団から得られる市販後調査報告からの副作用情報の解析が有用である。計算機科学的・統計学的解析手法は、創薬段階で得られた知見とともに、このような大規模で異なる属性を有するデータを網羅的に解析する力強い手段を提供する。このことを背景として、本研究は計算機による副作用発現機序推定を目的とした次の研究I、IIから構成される。</p> <p>研究Iでは、薬物と標的タンパク質の相互作用情報を用いることが副作用発現の機序推定を可能にするのではないかと立場から解析し、その結果を副作用予測へと応用した。685個の薬物について、医薬品関連データベースに登録された情報に基づき、1,368個の標的タンパク質への作用を記述したプロファイルと1,339個の副作用の有無を記述したプロファイルを作成した。機械学習のひとつであるスパース正準相関分析を適用して、互いに相関のある薬物-標的タンパク質-副作用という3種類の属性からなるグループを80個同定した。次に、生命知識データベースに登録されたタンパク質の分子経路情報と分子機能分類情報を用いてこれらのグループを特徴付けた結果、化学構造の異なる薬物が標的タンパク質を共有しないにも関わらず類似した副作用を発現するケースには、分子経路全体への制御が働いていると推定された。本解析手法を副作用の予測に応用したところ、既存の化学構造情報に基づく予測方法を上回る精度を達成した。また、副作用未知の薬物への予測に対しても文献による妥当性を確認した。</p> <p>研究IIでは、大規模市販後報告データに基づき、副作用の背景となる共通の仕組みを検出することを目的とした。FDA Adverse Event Reporting Systems (FAERS) から市販後有害事象報告データを取得し、1,317個の薬物と1,374個の副作用について692,707件の有害事象(425,273報告)からなる副作用の頻度行列を作成した。この行列にデータマイニングの方法であるバイクラスタリングを適用して163個の副作用クラスタを同定した。各クラスタには、発生頻度パターンの類似した副作用の集合と、それに対応する薬物の集合が抽出された。医学用語の階層的分類システムと、薬物の適応症データを用いて各クラスタを評価した結果、163個のうち145個のクラスタでは有害事象が薬物の適応症に非依存性である可能性が高いことが示された。従ってこの145個の各クラスタに抽出された副作用には、適応症以外に共通の薬理作用や患者の背景が関与していると推察された。そこで各クラスタに注目し、適応症の異なる薬物から副作用が起こる背景をマニュアルで精査した結果、共通の薬理作用が関与している例を見いだした。また、クラスタ間で副作用の種類を網羅的に比較した結果、これらが同一の薬物から起こる背景に患者の性別、年齢、体重の違いが統計的に有意に関与している例を検出した。これらの結果は、副作用の種類に共通の薬理作用や患者の特定の生理学的背景が関与していることを示唆している。</p> <p>本解析手法は、副作用の発現機序や背景を推定する手段として、計算機科学的・統計学的解析方法に基づいた医薬品関連データと生命知識の統合活用が有効であること示唆するものである。</p>			

(論文審査の結果の要旨)

申請者は、医薬品の副作用の発現機序を推定するために、医薬品標的タンパク質が関わる分子間相互作用ネットワークや医薬品の市販後調査報告という大規模データを有効に活用するユニークな研究を行った。論文では、まず薬物と標的タンパク質との相互作用情報を副作用と関連付ける新しい方法を開発し、その成果を薬物の副作用予測に応用した。続いて、薬物の大規模市販後報告データに基づき、副作用の背景となる共通の仕組みを検出する方法を提案して、患者の体重、年齢、性別との関係について解析した。

標的タンパク質情報を用いた副作用の解析では、スパース正準相関分析を応用して、互いに相関のある薬物、標的タンパク質、副作用からなる 80 個のグループを抽出した。この解析では、標的タンパク質と副作用がそれぞれ少なくとも一つは明らかになっている 685 個の薬物を用いて、1,386 個の標的タンパク質の対応関係と、同じ 685 個の薬物に対する 1,339 種類の副作用の対応関係をを入力とした。抽出された各グループにおける標的タンパク質を KEGG と Gene Ontology を用いて特徴付けたところ、同じ分子機能に分類されるものよりも同じ分子経路で働くものの方が多いことが明らかになった。このことは、異なる標的タンパク質が共通の副作用を発現する仕組みの一端を示している。また、本解析手法を応用して、標的タンパク質を元に副作用を予測したところ、薬物の化学構造を元に予測する方法よりもよい精度を示した。さらに、副作用未知の薬物に対する副作用予測結果についても、文献調査により妥当性も確認した。

薬物の大規模市販後報告データを用いた後半の研究では、米国 FDA が提供する市販後報告データから 1,317 個の薬物と 1,375 種類の副作用からなる頻度行列を構築し、バイクラスタリングを適用した。その結果、互いに相関する薬物と副作用のクラスタを 163 個抽出できた。抽出された各クラスタにおける薬物と副作用を適応症と副作用の類似度で特徴付けたところ、163 個のうち 145 個のクラスタで、副作用が適応症に非依存적であることが判明した。さらに、異なるクラスタに分類された共通の薬物について、クラスタ間の違いを検討した結果、同一の薬物から異なる副作用が発現する背景に患者の性別、年齢、体重の違いが統計的に有意に関与している例を検出できた。

本研究で抽出されたクラスタは、これまでほとんどブラックボックスであった、副作用の発現機序について、分子的な機序を解明するための大きな手がかりになると考えられる。また、本研究で開発・応用した手法は副作用の分子レベルでの発現機序解明に役立つだけでなく、副作用の予測への大きな手がかりになると考えられる。よって、本論文は博士（理学）の学位論文として価値あるものと認める。また、平成 27 年 1 月 14 日に論文内容とそれに関連した口頭試問を行った。その結果、合格と認めた。